

Blended Learning based on H5P Interactive Exercises: Insights from a Case Study

Michel Noutcha^[0009–0008–7103–0297] and Suzanne Kieffer^[0000–0002–5519–8814]

Université catholique de Louvain, Institute for Language and Communication,
Louvain-la-Neuve, Belgium
`{michel.noutcha,suzanne.kieffer}@uclouvain.be`

Abstract. The user experience (UX) in educational settings is crucial, extending beyond academic performance to encompass students’ perceptions and responses to teaching methods. Blended learning (BL), integrating online and face-to-face components, including flipped learning (FL), enhances engagement through platforms like the HTML5 package (H5P). Despite these advancements, guidelines for designing and evaluating online learning experiences with H5P are lacking. This paper aims to address this gap by providing recommendations for H5P integration and a methodology for assessing student learning experiences. Key contributions include positioning H5P among comparable tools, offering a detailed process for leveraging H5P, and emphasizing the need for effective assessment methods. The case study, conducted in a master’s thesis methodological seminar, involved two rounds of user testing, revealing insights into student preferences and needs. Findings suggest students prioritize content and usability over attractiveness and value H5P interactive features. Future work should replicate this case study with other teaching units, e-learning technologies, or other student groups.

Keywords: Blended learning · H5P · User Experience · Learning experience · Case study · Recommendations

1 Introduction

User experience (UX) encompasses individuals’ perceptions and responses resulting from using and/or anticipating the use of a product, system or service [18]. Beyond usability [21], UX requires examination of both pragmatic (e.g. ease of use) and hedonic attributes (e.g., stimulation) [13]. Aligned with UX, the learning experience extends beyond academic performance to include students’ perceptions and responses to an educational approach.

Blended learning (BL) merges online and face-to-face components [4, 17, 23], incorporating flipped learning (FL) where students access learning materials such as videos, tutorials, or documents online before attending classes. Plugins such as HTML5 Package (H5P)¹ enrich learning materials with interactive videos, fill-in-the-blanks or image sequencing, which empowers students to engage directly with educational content [5, 41, 36], enhancing the learning experience [7, 35].

¹ H5P:<https://h5p.org>

However, specific guidelines for the design and evaluation of H5P-based learning experiences are lacking, which raises questions such as: How to design online learning experiences that fully leverage the advantages of H5P? To what extent does the integration of H5P exercises into learning materials benefit students' overall learning experience? These are relevant questions to address given the ever-increasing demand for e-learning tools, especially since the successive lockdowns during the covid-19 pandemic. Studying the relevance of these tools is crucial for developing meaningful online learning experiences.

The objective of this paper is twofold: (1) provide recommendations for integrating H5P elements into pedagogical frameworks; (2) provide a methodology for assessing students' learning experience with H5P material in terms of performance, perceptions, and responses. The paper makes the following contributions:

1. The related work section positions H5P among comparable e-learning tools such as Xerte, Storyline, Adapt or eXeLearning, both informing current practices and establishing a foundation for potentially replicating this research, which focuses on H5P, with other e-learning tools.
2. The case study provides a detailed, step-by-step breakdown of how to leverage H5P interactive exercises to flip content. Special emphasis is placed on the critical role of designing H5P exercises that seamlessly integrate into pedagogical frameworks, thereby enhancing the effectiveness of learning.
3. The paper underscores the need for effective assessment methods and tools, using insights from a case study that evaluates students' performance and learning experiences with H5P exercises. It demonstrates the relevance of H5P exercises in fostering student engagement with the learning material.

The case study focuses on the integration of H5P exercises in the instructional design of a course taught at Université catholique de Louvain (UCLouvain). We integrated the H5P exercises in Moodle, an open-source learning management system (LMS) [1]. We conducted two rounds of user testing to evaluate students' academic performance and learning experience with the H5P material. A mixed-method approach allowed us to collect data from multiple data sources, cross-analyze these data, and report an integrated vision of the case. We adopted a case study approach, as scientific literature about the learning experience with H5P is scarce and we have no intent to statistically generalize findings or demonstrate cause-and-effect [29]. Instead, we discuss the proposed assessment methodology and recommendations, and the replication of the case study across other teaching units or e-learning technologies.

2 Background

2.1 Blended learning (BL)

BL, often used interchangeably with hybrid learning [17], lacks agreement on its definition in the scientific literature [40]. Generally, it refers to the integration of face-to-face (F2F) and online instruction [4, 17, 5]. Online learning occurs remotely via connected devices, either synchronously or asynchronously [8, 10, 30].

Two influential BL definitions are offered by Graham [11], emphasizing F2F with computer-mediated instruction, and Garrison [9], focusing on integrating F2F and online experiences. Graham’s definition, more inclusive, omits constraints on digital technology spaces, while Garrison’s emphasizes spatial, temporal, and human aspects of learning [8, 10].

Three meta-analyses consistently report that BL outperforms other teaching modalities, namely F2F [3, 39] and online [27, 39]. In particular, the authors show positive effects on knowledge outcome, academic performance, and enhancing education. These results can be explained as follows. From the instructor’s perspective, BL combines F2F and online modalities to address their respective shortcomings [7]. F2F lectures facilitate direct interaction, fostering a strong connection to the course structure and speakers, while online learning eliminates spatial and temporal barriers, reduces operational costs, and enhances instructors’ monitoring capabilities, allowing for personalized educational activities [24]. From the student’s perspective, BL cultivates 21st century skills such as creativity, critical thinking, and collaboration [19], while offering greater flexibility [22, 35] and student engagement [12].

However, BL remains challenging for both students and instructors. The main challenges for students are related to self-regulation and the use of educational technologies. As for instructors, the challenges are primarily linked to integrating technology for pedagogical purposes [31]. There are also security and privacy issues associated with using the H5P Hub to share and reuse interactive content, as well as physical and time constraints on students and instructors to access and use online resources effectively [5, 35].

2.2 Flipped learning (FL)

FL empowers students with more responsibility [14], by providing them with access to learning materials such as videos, tutorials, or documents online before F2F classes. Therefore, students’ initial exposure to new concepts or content occurs through self-directed online learning, allowing class time to be utilized for more interactive and collaborative activities, discussions, or application of knowledge. This approach contrasts with instructor-centered methods where students typically receive instructional content during class time and then work on assignments or activities independently outside of class.

Several tools exist to flip content (Table 1), among which HTML5 Package (H5P)². H5P is a free and open-source framework supporting the creation, sharing, and reuse of interactive HTML5 content. The framework includes an editor that allows creation and sharing H5P content, a website, and plugins for integration into learning management systems (LMS). The website contains examples of content types, a guide for authors and developers, and a forum for sharing libraries, applications, and content types. To date, H5P offers 52 content types among which quizzes, puzzles, interactive videos, and branching scenarios.

² H5P:<https://h5p.org>

Table 1. Comparison between H5P, Xerte, Storyline, Adapt, eXeLearning. Analysis conducted in Jan 2024. y stands for yes.

Systems	H5P	Xerte	Storyline	Adapt	eXeLearning
Last version	2024	2023	2023	2019	2013
Open-source	y	y		y	y
SCORM & HTML5	y	y	y	y	y
Other standards	y	y	y	y	limited
Reusable content	y	y	y	y	y
Community support	y	y	y	y	y
Responsiveness	y	y	y		y

In exploring H5P applications in education, we first delved into its positive impacts on student engagement, motivation, and self-regulation. For example, H5P branching scenarios enhance the learning experience within virtual simulations [35], while H5P interactive videos facilitate active engagement, self-paced review, and immediate feedback [22]. Additionally, H5P interactive activities contribute to asynchronous instructor presence [36]. Our exploration then extended to studies leveraging H5P to enhance video-based flipped classes [23, 41] or exploring various content types for out-of-class activities, such as branching scenario [37], accordion, image hotspot, and image sequencing [36]. Importantly, none of these works uses the 'column' content type, an H5P feature allowing the grouping of various H5P content types in a column layout. Our analysis reveals an untapped potential in utilizing this 'column' feature for a more integrated and dynamic learning experience. Moreover, to our knowledge, none explicitly investigates the UX with H5P or provides recommendations for flipping a course.

3 Case study

3.1 Context

We developed H5P resources for the master's thesis methodological seminar in the communication sciences program at UCLouvain. The seminar aims to equip students with essential research skills, including defining a research topic, anticipating methodological developments, evaluating project feasibility, and planning the research steps and formal requirements for completing the master's thesis. The 120-unit, two-year master program involves the seminar (2 units), pre-thesis research (3 units), and a master's thesis (20 units). The seminar assignment focuses on synthesizing future research elements, while pre-thesis research requires a comprehensive paper covering problem statement, literature review, methods, initial data, and research plan.

The shift to FL was prompted in 2018 by the combination of the following observations: students' stress and frustration due to the lack of individual follow-up; poor quality of the research topics submitted by the students at the end of the seminar; course holder's work overload while answering students' individual requests and grading students' research topic. In response, academic staff

decided to flip the class about bibliography management, and initiated group tutoring during the seminar. Group tutoring reduced both students' stress and course holder's work overload, and increased the quality of the research topics proposed by students. The resulting instructional design includes F2F lectures and discussions, individual assignments, in-class quizzes, and H5P exercises accessible on Moodle.

3.2 Flipping content with H5P

We flipped bibliography management for several reasons. First, it is universally applicable across all research topics, catering to the needs of all students. Second, mastering these activities is crucial for scientific and master's thesis writing, offering valuable refreshers for students. Last, flipped content replaces less engaging readings and lectures, optimizing F2F time for more pertinent subjects.

Although BL is often associated with video-based classes, we diverged from the choice of H5P interactive video, the most extensively discussed H5P content type in scientific literature. Instead, we opted for the H5P column for the following three reasons. First, relying solely on interactive videos for FL would have been time-consuming, since video preparation and maintenance is resource-intensive [41]. Second, interactive videos, proven to be most effective within 5-10 minutes [23, 35], could risk losing student interest when dealing with a larger quantity of content to flip. Third, by utilizing H5P columns, we could consolidate the content associated with a specific teaching unit within a single column, adhering to Mayer's segmenting principle [26]. Specifically, we chunked the content into four chapters, database searching (Chap.1), bibliography (Chap.2), referencing (Chap.3) and avoiding plagiarism (Chap.4). We created a fifth H5P column to introduce students to H5P interaction. For each H5P exercise (i.e., chapter), we followed the same procedure: (1) document the content of the corresponding teaching unit; (2) assign H5P modalities to the written content; (3) implement in H5P; (4) proof-read and check the result as validation.

Table 2. Description of the H5P content used in each H5P column/exercise.

	Chap.1	Chap.2	Chap.3	Chap.4
https://h5p.org/drag-and-drop	2	11	0	3
https://h5p.org/drag-the-words	4	5	1	2
https://h5p.org/fill-in-the-blanks	0	3	0	0
https://h5p.org/image-hotspot-question	0	4	0	0
https://h5p.org/image-hotspots	1	3	0	0
https://h5p.org/image-slider	0	5	0	0
https://h5p.org/single-choice-set	2	0	4	3
https://h5p.org/true-false	0	3	1	3
https://h5p.org/interactive-video	2	4	2	2

3.3 User testing

We employed user testing to gather data allowing us to assess participants' UX, academic performance, and attitude toward the H5P exercises. User testing involved questionnaire to measure UX, knowledge test to measure academic performance, and interview or questionnaire to gather data about attitude. This mixed-method approach allowed us to cross-analyze different types of data.

We proceeded in two rounds. Round 1 (R1) was exploratory and took place in distant synchronous mode in June 2021 with $N = 12$ participants, while Round 2 (R2) was descriptive and took place in online asynchronous mode in May 2023 with $N = 27$ participants. Both rounds followed a similar procedure. First, all participants signed an online consent form authorizing the use of their anonymized data for research purposes. Following this, they performed the experimental tasks. Finally, they filled out a questionnaire regarding UX, which was followed by either an interview (R1) or a questionnaire (R2) about their attitude toward the H5P material. There were differences between R1 and R2 regarding the user testing mode, the scope of the experimental task, the pretest questionnaire, and the knowledge tests (Table 3). First, R1 took place synchronously on Teams using links to the different instruments (e.g. consent form, tests, etc.) with a video recording of the participant's screen shared while they performed the experimental task on Moodle. R2 took place completely asynchronously on Moodle. Second, R1 participants completed one or two H5P chapters, while R2 participants completed all four H5P chapters. Third, R2 participants completed a questionnaire assessing their level of expertise with each teaching unit and their level of familiarity with H5P. Finally, R2 participants underwent a pretest and posttest knowledge test, respectively before and after the completion of the experimental task.

Table 3. Overview of the methodology in both rounds. * indicates R1 activities accessed through a link provided by the experimenter in Teams conversation.

Round	R1	R2
Period	June 2021	May 2023
Goal	exploratory	descriptive
Mode	online synchronous	online asynchronous
Participants	$N = 12$	$N = 27$
Online consent form	yes*	yes
Pretest questionnaire	no	yes
Pretest knowledge test	no	yes
Experimental task	1 or 2 chapters	all 4 chapters
Posttest knowledge test	no	yes
UX questionnaire	yes*	yes
Follow-up	interview	questionnaire

Please only check one circle in each line.

	1	2	3	4	5	6	7	
Attractiveness								
In my opinion, the product is generally								
annoying	☐	☐	☐	☐	☐	☐	☐	enjoyable
bad	☐	☐	☐	☐	☐	☐	☐	good
unpleasant	☐	☐	☐	☐	☐	☐	☐	pleasant
unfriendly	☐	☐	☐	☐	☐	☐	☐	friendly
I consider the product property described by these terms as								
Completely irrelevant	☐	☐	☐	☐	☐	☐	☐	Very important

Fig. 1. Attractiveness scale. Template from <https://ueqplus.ueq-research.org/>

3.4 Data collection methods

User experience. To assess participants' UX with the H5P material, we used the UEQ+, a modular and standardized questionnaire for measuring UX according to up to 26 scales [33]. We used the same nine scales in both rounds. Specifically, we opted for efficiency, usefulness, perspicuity, and intuitive use, as these are recommended by authors for assessing systems designed to help users achieve specific goals. We also included trustworthiness and quality of content, as suggested for products conveying knowledge. We selected attractiveness, stimulation, and novelty as they apply to all products. We chose these scales due to their balanced coverage of task-oriented (e.g. usefulness) and non-task-oriented (e.g. stimulation) attributes of UX [13]. Each scale gathers ratings on a 1 to 7 scale for four related items and the relative importance of the scale (Figure 1). The ratings for the four related items collected from all participants enable the calculation of a mean UX score for each scale and an overall UX score for the H5P exercises. The mean relative importance of each scale is computed as the average of the importance ratings across all participants.

Academic performance. To assess R2 participants' academic performance (not assessed in R1), we administered Moodle knowledge tests. A pretest knowledge test measured participants' baseline understanding of each teaching unit, serving as a benchmark for evaluating teaching effectiveness. We then compared the posttest knowledge test results to assess participant progress and the efficacy of H5P exercises. Finally, we cross-analyzed data with participants' scores in each H5P exercise, which were accessible from Moodle gradebook. We developed a Moodle question database consisting of 20 questions in total, with five questions allocated to each H5P chapter. Our design prioritized competencies over knowledge assessment, with four questions on knowledge (i.e., what students know) and 16 questions focusing on competencies (i.e., what students can do with that knowledge) assessment, respectively. Examples of questions are provided in Table 4. Questions were randomly extracted from the Moodle question database for both the pretest and posttest knowledge assessments.

Table 4. Examples of R2 knowledge test questions by question type. Q: questions; A: answers; correct answers in bold; blank indicates no question for that question type.

Knowledge	Competency
Chapter 1: Database searching	
Q: What does bibliographic research entail? A: identifying previous research conducted on a specific topic ; referencing existing research on the topic; summarizing prior research findings on the topic	Q: What factors should be prioritized in analyzing the relevance of a document found in a scientific article database? A: abstract ; authors; page count; source
Chapter 2: Bibliography	
	Q: What type of document does "Lallemand, C., Gronier, G., & Dugué, M. (2018). Méthodes de design UX: 30 méthodes fondamentales pour concevoir des expériences optimales (2e éd). Eyrolles." correspond to? A: article; book ; book chapter; DVD
Chapter 3: Referencing	
Q: Where should all references be found? A: at the end of the text, in footnotes, in the list of references	Q: How to cite a source from a single author in parenthesis format? A: (author's name, publication date); (author's name, publication date, page); (publication date, author's name)
Chapter 4: Avoiding plagiarism	
Q: Why cite sources in academic work? (select all that apply) A: to avoid plagiarism ; ; to distinguish personal opinions from others' ; to enhance credibility ; to provide readers with additional reference	Q: How to avoid plagiarism when copy-pasting a verbatim excerpt into your own work? A: italicizing; italicizing, using quotation marks and referencing the source ; using quotation marks

Attitude. To assess R1 participants' attitudes toward H5P exercises, we conducted semi-structured interviews using the following questions: "Which aspects of the H5P exercises positively affected your satisfaction?", "What are the three most positive/negative aspects of the exercises?", and "Which aspects should we improve in the H5P exercises?". Additional interventions, such as "Could you tell us more about [...]?", "Do you have a concrete example of [...]?", or "What does [...] mean to you?", were used when needed to clarify participants' statements. For R2 participants' attitudes toward the H5P-based instructional design, we administered a questionnaire with the following questions: "What do you consider to be the positive aspects of this learning tool?", "What aspects of this learning tool do you believe could be improved?", and "Is there anything else you would like to add?". This approach allowed us to gather comprehensive insights into participants' perceptions and preferences regarding the H5P exercises and instructional design, facilitating informed decision-making for future improvements and enhancements.

3.5 Participants

We recruited all participants from the communication science master’s program at UCLouvain, thereby the sample is representative of the targeted user population [20]. Specifically, R1 participants were recruited through the social media of the master’s program, while R2 participants were all enrolled and active students in the UX course of the master’s program. Analysis of pretest questionnaire revealed varying levels of familiarity with H5P: in R1, 9 out of 12 participants (75%) had never used H5P before the user tests, whereas all R2 participants (100%) had used it at least once.

Proficiency in bibliography management showed consistent patterns among R2 participants (Table 5). Notably, they reported above average proficiency in all teaching units with slight variability ($Mean = 2.77$, $Median = 3$, $SD = .67$). Additionally, they reported a higher proficiency ($Mean = 2.96$) in avoiding plagiarism, presumably due to the severe consequences it carries in academia, while reporting lower proficiency in both database searching and bibliography ($Mean = 2.70$) with a smaller variability ($SD = .67$), suggesting a potential lack of confidence in their computational skills when searching databases and a potential lack of familiarity with tools for automated bibliography creation.

Table 5. Summary of R2 participants’ proficiency with chapters. Proficiency expressed on a 4-level scale from not proficient (1) to proficient (4). SD: standard deviation.

Proficiency	Mean	Median	SD
Database searching (Chap.1)	2.70	3	.67
Bibliography (Chap.2)	2.70	3	.67
Referencing (Chap.3)	2.81	3	.62
Avoiding plagiarism (Chap.4)	2.96	3	.71

3.6 Conducting user tests

Completing all four chapters typically takes around two hours. Therefore, to ensure that user tests remained within a reasonable duration of under 60 minutes, we randomized the assignment of R1 participants to different chapters. Participants P1, P4, P7, and P10 completed Chap.1; participants P2, P5, P8, and P11 completed Chap.2; and participants P3, P6, P9, and P12 completed both Chap.3 and Chap.4. This allocation was based on the number of H5P contents in each chapter (Table 2). We used random assignment to mitigate selection bias [38].

In R2, the user tests were intentionally conducted asynchronously on Moodle, closely mimicking the intended instructional setting. This allowed R2 participants to complete exercises at their own pace, without an experimenter present, promoting greater autonomy and enhancing action fidelity [20]. While R1 participants’ behavior may not correspond to the intended behavior due to artificial conditions, R2 participants’ behavior may align closely with the behavior expected in the natural instructional context.

We employed conditional access on Moodle to ensure that certain activities remained inaccessible until prerequisites were completed. As depicted in Figure 2, R2 participants were required to sign the consent form before proceeding to the pretest questionnaire and knowledge test, and to complete pretest activities before accessing the posttest knowledge test and questionnaires. We opted not to implement conditional access to the chapters to maintain action fidelity. This allowed participants the flexibility to skip chapters or complete them in any order they preferred. This choice was also driven by occasional issues where H5P activities are not always recognized as completed in Moodle, despite being so, making it challenging to assess completion.

To collect complete answers, the user test was treated as an assignment in the UX course, contributing to R2 participants' final grade. The user test would not be considered accomplished unless the last questionnaire was answered. The user test had a designated start and end date. We provided an estimated duration to complete each chapter.

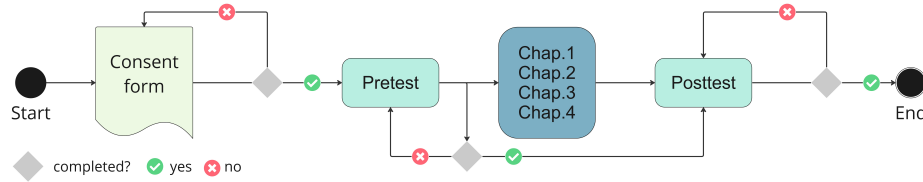


Fig. 2. Conditional access to Moodle activities.

4 Results

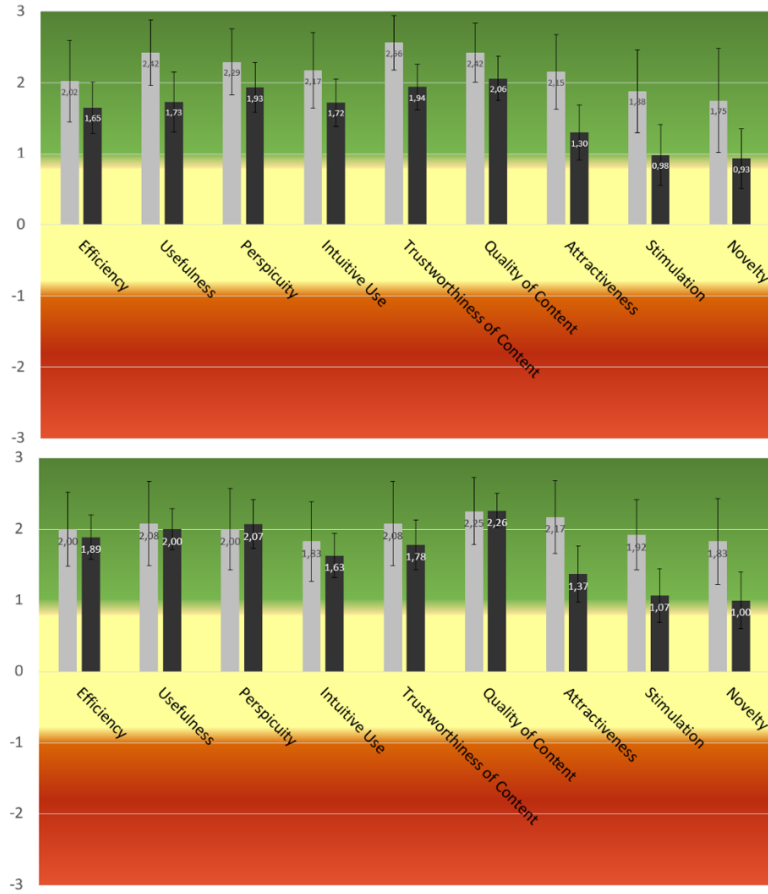
We collected both quantitative and qualitative data. R1 being exploratory and R2 descriptive, we limited quantitative data analysis to descriptive statistics. We used content analysis to analyze qualitative data, which consisted in transcribing answers, coding information in an analysis grid, and processing it.

4.1 User experience

We analyzed participants' UX with H5P exercises by looking at the data collected with the UEQ+. Table 6 presents UX scores for each scale, while Figure 3 depicts UX scores and relative importance. Altogether, UX scored very high in R1 ($Mean = 2.18$) and high in R2 ($Mean = 1.58$) indicating an overall very good UX with the H5P exercises. However, we observe a consistent decrease in UX scores across all scales between R1 and R2 (mean: $-.60$, range: $-.90$ to $-.36$), likely due to R2 participants' greater familiarity with H5P exercises and their acquired UX expertise from attending UX courses in contrast to R1 participants, and the much longer duration of user tests in R2.

Table 6. UX scores per round. Descending order on R2 means. SD: standard deviation.

UX scale	Mean		SD	
	R1	R2	R1	R2
Efficiency	2.02	1.65	1.01	0.96
Usefulness	2.42	1.73	0.81	1.1
Perspicuity	2.29	1.93	0.82	0.94
Intuitive use	2.17	1.72	0.94	0.88
Trustworthiness of content	2.56	1.94	0.67	0.84
Quality of content	2.42	2.06	0.73	0.81
Attractiveness	2.15	1.3	0.92	1.03
Stimulation	1.88	0.98	1.03	1.14
Novelty	1.75	0.93	1.28	1.1

**Fig. 3.** UX scores (top) and relative importance (bottom). Mean scores with confidence intervals per UX scale on a -3 to +3 range. R1 in light grey, R2 is in dark grey.

UX scores for efficiency, usefulness, perspicuity, and intuitive use are well above average in both rounds (consolidated means: 2.23 in R1, 1.76 in R2), highlighting the effectiveness of the H5P exercises in facilitating goal achievement. In particular, perspicuity scores indicate that learning how to use H5P was easy. In addition, participants in both rounds consistently expressed the high importance of these four UX attributes, especially in R2 compared to attractiveness, stimulation, or novelty.

Throughout both rounds, participants consistently expressed high regard for the trustworthiness and quality of content (consolidated means: 2.49 in R1, 2.00 in R2), as reflected in the respective mean UX scores with significant relative importance. This indicates that the information conveyed through the H5P exercises is reliable, of high quality, up-to-date, and well-prepared. Notably, the relative importance of quality of content did not decrease between rounds, which indicates that participants placed more value on substance than on other system characteristics, such as attractiveness, stimulation, or novelty.

UX scores for attractiveness, stimulation, and novelty exhibit the most notable decrease between R1 and R2 (consolidated means: 1.93 in R1, 1.07 in R2), indicating a significant decline in overall impression regarding the exercises' attractiveness, originality, and enjoyment. This decline could be attributed to R2 participants' increased familiarity with H5P exercises and their acquired UX expertise from attending UX courses, as well as the substantially longer duration of user tests in R2. Unsurprisingly, R2 participants also rated these system characteristics as less important compared to others. Although results demonstrate the usability of the H5P exercises as well as the validity of prioritizing content quality, they suggest that the entire H5P set could benefit from shortening the overall duration to complete the exercises, so as to increase attractiveness.

4.2 Academic performance

As there was no knowledge test in R1, the entire remainder of this section focuses solely on R2 and its participants. We analyzed R2 participants' academic performance by looking at the scores they obtained at the pre- and posttest knowledge tests (12 questions, 1 point each) and H5P exercises (10 points each). Participants exhibited improved performance in the posttest ($Mean = 10.21$) compared to the pretest ($Mean = 8.47$), with most scores clustered near the maximum ($Median = 10.75$) and a smaller range of scores from 6 to 12 (Table 7), which highlights the effectiveness of H5P exercises in improving knowledge and competencies regarding bibliography management.

Table 7. Comparison of academic performance between pretest and posttest.

Test	Mean	Median	Range	N
Pretest	8.47	8	3.85 - 11.5	27
Posttest	10.21	10.75	6 - 12	27

Most participants ($N = 18$, 67%) improved their scores between the pretest and posttest, two (7%) maintained identical scores, seven (26%) experienced decline (Table 8). The group that showed improvement increased performance by 26% (from 62% to 88%), while the group that experienced decline decreased performance by 10% (from 88% to 78%). This suggests H5P exercises are most beneficial for participants who demonstrated the least proficiency in the pretest.

A detailed examination of participants performance and engagement in H5P activities highlights a consistent trend of high scores across all chapters, with the majority of participants achieving 100% together with a strong engagement in Chap.1 ($N = 25$) and Chap.3 ($N = 23$). Chap.2 and Chap.4 appear less explored, with a majority of participants (respectively $N = 18$ and $N = 20$) failing to engage with them.

Table 8. Participants scores per chapter, pretest and posttest. Delta: score difference between posttest and pretest.

Participant	Chap.1	Chap.2	Chap.3	Chap.4	Pretest	Posttest	Delta
P8	10	-	10	-	5.5	11.8	6.30
P6	10	-	10	-	7.5	12	4.50
P20	10	10	10	8.89	7.5	12	4.50
P4	10	10	10	-	7.75	12	4.25
P26	10	-	10	-	8	12	4.00
P25	10	10	10	9.72	7	10.75	3.75
P11	10	-	10	-	7.5	11	3.50
P19	10	-	-	-	7.6	11	3.40
P14	10	-	10	-	7	10	3.00
P1	7.65	-	9.38	6.94	7.5	10.5	3.00
P21	10	-	9.38	-	6.5	9.35	2.85
P27	-	-	-	-	5.5	8	2.50
P3	10	10	10	9.72	9.5	12	2.50
P24	10	-	10	-	3.85	6	2.15
P2	10	10	10	9.72	10	12	2.00
P9	10	-	-	-	9.35	11.1	1.75
P23	10	-	10	-	6	7.25	1.25
P12	10	-	10	-	10	10.6	0.60
P7	10	-	10	-	10	10	0.00
P13	10	-	10	-	11	11	0.00
P5	10	10	10	8.89	11.25	11	-0.25
P16	10	10	10	-	11.50	11	-0.50
P17	10	10	10	10	10	9	-1.00
P22	10	10	8.75	-	11	10	-1.00
P10	6.47	-	8.75	-	10.25	8.50	-1.75
P15	-	-	10	-	10.25	8.25	-2.00
P18	3.53	-	-	-	10	7.75	-2.25
Mean	9.50	10.00	9.84	9.13	8.47	10.21	1.74
N	25	9	23	7	27	27	27

Notably, 88% of participants achieved a score of 10 in Chap.1, 100% in Chap.2, 82.60% in Chap.3, and only 14% in Chap.4, demonstrating a solid grasp of the material except in Chap.4. Participants exhibited a strong engagement in Chap.1 (92.59%) and Chap.3 (85.19%), indicating interest in the themes, likely because database searching and referencing are areas where they typically encounter the most difficulty. Conversely, participants shown disinterest in both Chap.2 and Chap.4, likely overestimating their knowledge and competencies in bibliography and plagiarism.

4.3 Attitudes

Thematic analysis of participants feedback reveals recurring themes, providing insight into positive aspects and areas for improvement in the H5P material. Positive aspects include the richness of H5P contents, their interactive features, and their ability to empower participants. First, participants praised the diversity of content types offering a good user experience, ranging from drag-and-drop exercises to interactive videos. In particular, participants found the combination of video and text beneficial as it enabled them to easily follow along with the text while watching the video (P12 in R1, P25 in R2). In R2, P8 mentioned that the variety of questions helped in retaining the information read. Second, participants enthusiastically welcomed interactive features such as drag-and-drop and interactive videos, despite manipulation slips while placing tiles in drag-and-drop (P9 in R1, P6 and P13 in R2) and some activities exceeding the page size (P6 and P20 in R2). However, interactions kept them engaged (P6 in R1, P2 in R2), focused (P8 in R2), while making exercises enjoyable to use (P5 in R2). Third, participants valued the ability to self-correct thanks to direct feedback after answers, the flexibility to study at their own pace, the ability to access exercises at any time (e.g. P6 in R1 and P8 in R2). Furthermore, P2 in R1 and P25 in R2 judged the possibility to retry as motivating. These results are consistent with the relevant literature regarding testing effect [36], interactivity [22], stimulation [37], and interactive videos [41].

Suggestions for enhancement mainly centered around augmenting the H5P column with new features, such as a progress bar (P8 in R1, P6 in R2), which is beyond our control. P11 in R2 also suggested incorporating accessibility features for students with special needs. Additionally, participants expressed concerns about the excessive length of Chap.2. Lastly, two participants (P1 in R1, P23 in R2) highlighted the need for additional feedback in H5P activities, regardless of whether they provided correct answers or not. This issue is beyond our control since feedback in H5P is determined by a range of correctness rather than individual items. However, we intend to meet this request in a future version of Moodle pre- and posttest knowledge tests.

5 Discussion

5.1 Cross-analyzing data

Participants from R1 and R2 exhibited a high level of satisfaction with the H5P exercises, as evidenced by both UX scores and attitudes. Their satisfaction was particularly pronounced regarding content (trustworthiness, quality, and richness) and they perceived the H5P exercises as beneficial to learning. While usability factors (efficiency, usefulness, perspicuity, and intuitive use) also contributed to their satisfaction, they received less emphasis, especially efficiency and intuitive use. This is likely due to minor difficulties (e.g. placing tiles in drag-and-drop features and content exceeding page size), which required participants to redo their work to achieve their goals. These minor shortcomings are largely balanced out by the fact that H5P interactive features kept participants motivated and focused on the content. These findings demonstrate that the information conveyed in H5P exercises is reliable, well-crafted, and engaging, which is crucial for the effectiveness of online learning [39].

The pronounced decrease between R1 and R2 in attractiveness, stimulation, and novelty UX scores can be attributed to two factors: increased familiarity of R2 participants with H5P and excessive duration of Chap.2. The duration of Chap.2 also elucidates the disinterest of R2 participants, as evidenced by their lack of engagement with it. The comparable and simultaneous decrease in relative importance suggests that participants prioritized substance over form, viewing H5P as a task-oriented technology in learning contexts. Task-oriented refers to technologies that users employ to achieve specific goals (here, learning), whereas non-task-oriented refers to technologies where users formulate goals "on the fly" [13]. This finding underscores the importance of crafting H5P exercises that capture students' interest and prompt them to fully engage with them [12].

Moreover, the improvement in academic performance observed between pre- and posttest knowledge tests suggests that the H5P exercises effectively facilitate knowledge and competencies acquisition. This finding corroborates existing literature indicating the efficacy of H5P in promoting deeper understanding and mastery of subject matter [5, 22, 35]. Chap.1 and Chap.3, which elicited higher levels of engagement, exhibited superior academic performance, highlighting the importance of designing engaging and relevant learning materials. The association between R2 participants' engagement and academic performance further supports the notion that active involvement in learning activities contributes to better learning outcomes [12].

Finally, participants' feedback provided valuable insights into the strengths and areas for improvement of the H5P exercises. Positive feedback regarding the variety of content types and user-friendly interactive features underscores the importance of providing diverse and engaging learning experiences [6]. However, suggestions for enhancements, such as accessibility features for students with special needs, indicate areas where further development is needed to meet the evolving needs of users.

5.2 Strengths and limitations

The improvement in academic performance observed between pre- and posttest knowledge tests supports our initial objective of flipping content to allocate more F2F lecture time to other relevant topics, while addressing the methodological needs of all students regarding a core aspect of completing a master’s thesis. By equipping students with a solid foundation in bibliographic management, H5P exercises have brought us closer to achieving this goal.

We did not intend to make statistical generalizations. Case studies typically intend to provide an in-depth understanding of a specific case. Instead, we report an integrated vision of multiple qualitative and quantitative variables regarding students’ user experience, academic performance, and attitude toward learning material flipped with H5P. R2 assessment methodology can be replicated with other courses, curricula, and other e-learning tools, such as Xerte, Storyline, Adapt or eXeLearning. Cross-analysis between multiple variables with $N = 39$ participants is a strength rather than a weakness. However, it would be interesting to replicate the case study with student groups in other master’s programs.

6 Recommendations

BL requires changing the habits and practices of both students and instructors [34]. Specifically, instructors may face challenges while flipping content for BL. The following recommendations are intended for instructors willing to use H5P to flip content and range from practical tips (1-4) to conceptual advice (5-8), positioned against related work.

1. Start small with existing unflipped resources. The implementation of H5P exercises is resource-intensive [41]: implementing our first H5P exercise took us a hundred times the student time to complete the exercise. We recommend starting with a small chunk [26] of course that will produce a substantial return on investment, e.g. a refresher course.
2. Use one H5P column per teaching unit. Instructors can consolidate content for a teaching unit in one H5P column, aligning with Mayer’s segmenting principle [26]. Structuring these elements in a column layout introduces a visually engaging dimension, promoting clear comprehension [26]. This versatile approach optimizes student engagement, elevating the learning experience’s overall effectiveness.
3. Be economical with both content and instructions. Superfluous elements can impede learning by requiring excessive cognitive processing, hindering educational objectives [6, 26]. Additionally, maintaining video duration of around five minutes is advised to mitigate dropout rates [23, 35]. Finally, avoid spending time producing a getting started tutorial, as H5P is intuitive and easy to use.
4. Allow students to manage the pace of their learning by indicating time to complete exercises and providing unconstrained access to material. This promotes self-regulation by creating an opportunity for students to learn how to control their time and to complete exercises in the sequence they choose.

5. Establish instructor-student relationship through feedback in answers [25, 32]. Although this recommendation requires a significant resource investment, it is worth the effort, as it is important to provide feedback to guide and motivate underperforming students and to mark the presence of the physically absent instructor.
6. Allow for game-based learning [28, 2] through various H5P contents ranging from well-known interactions such as drag-and-drop to lesser-known such as image hotspots. The playfulness of H5P interactions (e.g. earning points and stars) will help maintain students' motivation [35, 22].
7. Promote inclusive and accessible teaching by using multiple modalities (e.g. text, images, audio, video) to present the same information. This "efficient redundancy" addresses the specific needs of a wide range of students with diverse abilities and learning styles [6, 26].
8. Flip common-ground content that will meet a need encountered by all students. We recommend selecting refresher courses, or a document or a lecture that students typically consider boring, to allocate more F2F lecture time to advanced topics and students' requests.

7 Conclusions

In this paper, we adopted a user-centered approach to flip course content using H5P and delivered an instructional design that meets students' needs [42]. Further, we demonstrated that H5P exercises satisfy the increasing need for active pedagogy aiming at making students actors of their learning [15]. Through user testing of H5P material targeted at bibliography management, we demonstrated that H5P helps maintain students' motivation, and supports self-directed, game-based and trial-and-error learning, which enhances learning experiences [16]. However, ongoing efforts are required to address students' specific needs. The case study provides a detailed, step-by-step breakdown of how to leverage H5P interactive exercises to flip content, as well as a series of recommendations to flip content using H5P. Future work should replicate the case study with student groups in other master's programs.

Acknowledgments. The authors acknowledge the support provided by the LECaaS project (reference 8610), funded by Service public de Wallonie (SPW) and by the Institute for Language and Communication. Appreciation is also extended to the anonymous reviewers for their insightful feedback on an earlier iteration of this work. Lastly, the authors warmly thank all the participants in the user testing, and Mrs. Estelle Pepinster, whose significant contributions as part of her master's research thesis were instrumental in the design and implementation of H5P exercises.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Al-Ajlan, A., Zedan, H.: Why moodle. In: 2008 12th IEEE International Workshop on Future Trends of Distributed Computing Systems. pp. 58–64. IEEE (2008). <https://doi.org/10.1109/FTDCS.2008.22>
2. Alsawaier, R.S.: The effect of gamification on motivation and engagement. The International Journal of Information and Learning Technology **35**(1), 56–79 (2018). <https://doi.org/10.1108/IJILT-02-2017-0009>
3. Bernard, R.M., Borokhovski, E., Schmid, R.F., Tamim, R.M., Abrami, P.C.: A meta-analysis of blended learning and technology use in higher education: From the general to the applied. Journal of Computing in Higher Education **26**, 87–122 (2014). <https://doi.org/10.1007/s12528-013-9077-3>
4. Boelens, R., De Wever, B., Voet, M.: Four key challenges to the design of blended learning: A systematic literature review. Educational Research Review **22**, 1–18 (2017). <https://doi.org/10.1016/j.edurev.2017.06.001>
5. Chen, L., Manwaring, P., Zakaria, G., Wilkie, S., Loton, D.: Implementing h5p online interactive activities at scale. ASCILITE 2021: Back to the Future–ASCILITE ‘21 Proceedings ASCILITE 2021 in Armidale (2021). <https://doi.org/10.14742/ascilite2021.0112>
6. Clark, R.C., Mayer, R.E.: E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning. John Wiley & sons (2016)
7. Dakhi, O., JAMA, J., IRFAN, D., et al.: Blended learning: a 21st century learning model at college. International Journal Of Multi Science **1**(08), 50–65 (2020)
8. Dhawan, S.: Online learning: A panacea in the time of covid-19 crisis. Journal of educational technology systems **49**(1), 5–22 (2020). <https://doi.org/10.1177/0047239520934018>
9. Garrison, D.R., Kanuka, H.: Blended learning: Uncovering its transformative potential in higher education. The internet and higher education **7**(2), 95–105 (2004). <https://doi.org/10.1016/j.iheduc.2004.02.001>
10. George, P.P., Papachristou, N., Belisario, J.M., Wang, W., Wark, P.A., Cotic, Z., Rasmussen, K., Sluiter, R., Riboli-Sasco, E., Car, L.T., et al.: Online elearning for undergraduates in health professions: a systematic review of the impact on knowledge, skills, attitudes and satisfaction. Journal of global health **4**(1) (2014). <https://doi.org/10.7189/jogh.04.010406>
11. Graham, C.R.: Blended learning systems. The handbook of blended learning: Global perspectives, local designs **1**, 3–21 (2006)
12. Halverson, L.R., Graham, C.R.: Learner engagement in blended learning environments: A conceptual framework. Online Learning **23**(2), 145–178 (2019). <https://doi.org/10.24059/olj.v23i2.1481>
13. Hassenzahl, M.: The thing and i: understanding the relationship between user and product. Funology 2: from usability to enjoyment pp. 301–313 (2018). https://doi.org/10.1007/978-3-319-68213-6_19
14. Hew, K.F., Lo, C.K.: Flipped classroom improves student learning in health professions education: a meta-analysis. BMC medical education **18**, 1–12 (2018). <https://doi.org/10.1186/s12909-018-1144-z>
15. Hilliard, A., Kargbo, H.F.: Educationally game-based learning encourages learners to be actively engaged in their own learning. International Journal of education and Practice **5**(4), 45–60 (2017). <https://doi.org/10.18488/journal.61.2017.54.45.60>
16. Homanová, Z., Havlásková, T.: H5p interactive didactic tools in education. In: 11th International Conference on Education and New Learning Technologies. pp. 9266–9275 (2019)

17. Hrastinski, S.: What do we mean by blended learning? *TechTrends* **63**(5), 564–569 (2019). <https://doi.org/10.1007/s11528-019-00375-5>
18. ISO 9241-210:2019: Ergonomics of human-system interaction – Part 210: Human-centred design for interactive systems. Standard, International Organization for Standardization, Geneva, CH (2019)
19. Karaca-Atik, A., Meeuwisse, M., Gorgievski, M., Smeets, G.: Uncovering important 21st-century skills for sustainable career development of social sciences graduates: A systematic review. *Educational Research Review* p. 100528 (2023). <https://doi.org/10.1016/j.edurev.2023.100528>
20. Kieffer, S.: Ecoval: Ecological validity of cues and representative design in user experience evaluations. *AIS Transactions on Human-Computer Interaction* **9**(2), 149–172 (2017)
21. Kieffer, S., Rukonic, L., Kervyn de Meerendré, V., Vanderdonckt, J.: Specification of a ux process reference model towards the strategic planning of ux activities. In: *VISIGRAPP (2: HUCAPP)*. pp. 74–85 (2019). <https://doi.org/10.5220/0007693600740085>
22. Killam, L.A., Luctkar-Flude, M.: Virtual simulations to replace clinical hours in a family assessment course: Development using h5p, gamification, and student co-creation. *Clinical Simulation in Nursing* **57**, 59–65 (2021). <https://doi.org/https://doi.org/10.1016/j.ecns.2021.02.008>, <https://www.sciencedirect.com/science/article/pii/S1876139921000281>
23. Kuran, M.Ş., Pedersen, J.M., Elsner, R.: Learning management systems on blended learning courses: An experience-based observation. In: Choraś, M., Choraś, R.S. (eds.) *Image Processing and Communications Challenges 9*. pp. 141–148. Springer International Publishing, Cham (2018). https://doi.org/10.1007/978-3-319-68720-9_17
24. Lhafa, F.Z., Abdoun, O.: Adaptive collaborative learning process in a hybrid model. In: *Proceedings of the 8th International Conference on Advanced Intelligent Systems and Informatics 2022*. pp. 26–38. Springer (2022). https://doi.org/10.1007/978-3-031-20601-6_3
25. Lucas, G.M.: Initiating student-teacher contact via personalized responses to one-minute papers. *College Teaching* **58**(2), 39–42 (2010). <https://doi.org/10.1080/87567550903245631>
26. Mayer, R.E.: Using multimedia for e-learning. *Journal of computer assisted learning* **33**(5), 403–423 (2017). <https://doi.org/10.1111/jcal.12197>
27. Means, B., Toyama, Y., Murphy, R., Baki, M.: The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers college record* **115**(3), 1–47 (2013). <https://doi.org/10.1177/016146811311500307>
28. Padilla-Meléndez, A., del Aguila-Obra, A.R., Garrido-Moreno, A.: Perceived playfulness, gender differences and technology acceptance model in a blended learning scenario. *Computers & Education* **63**, 306–317 (2013). <https://doi.org/10.1016/j.compedu.2012.12.014>
29. Quintão, C., Andrade, P., Almeida, F.: How to improve the validity and reliability of a case study approach? *Journal of Interdisciplinary Studies in Education* **9**(2), 264–275 (2020). <https://doi.org/10.32674/jise.v9i2.2026>
30. Raes, A., Detienne, L., Windey, I., Depaepe, F.: A systematic literature review on synchronous hybrid learning: gaps identified. *Learning Environments Research* **23**, 269–290 (2020). <https://doi.org/10.1007/s10984-019-09303-z>
31. Rasheed, R.A., Kamsin, A., Abdullah, N.A.: Challenges in the online component of blended learning: A systematic review. *Computers & Education* **144**, 103701 (2020)

32. Reinholz, D.L., Dounas-Frazer, D.R.: Personalized instructor responses to guided student reflections: Analysis of two instructors' perspectives and practices. *American Journal of Physics* **85**(11), 850–860 (2017)
33. Schrepp, M., Thomaschewski, J.: Design and validation of a framework for the creation of user experience questionnaires. *International Journal of Interactive Multimedia & Artificial Intelligence* **5**(7) (2019). <https://doi.org/10.9781/ijimai.2019.06.006>
34. Senali, M.G., Iranmanesh, M., Ghobakhloo, M., Gengatharen, D., Tseng, M.L., Nilsashi, M.: Flipped classroom in business and entrepreneurship education: A systematic review and future research agenda. *The International Journal of Management Education* **20**(1), 100614 (2022). <https://doi.org/https://doi.org/10.1016/j.ijme.2022.100614>, <https://www.sciencedirect.com/science/article/pii/S1472811722000167>
35. Singh, S., Scholz, K.: Using an e-authoring tool (h5p) to support blended learning: Librarians' experience. In: *Me, Us, IT! Proceedings ASCILITE2017: 34th International Conference on Innovation, Practice and Research in the Use of Educational Technologies in Tertiary Education*. pp. 158–162 (2017)
36. Sinnayah, P., Salcedo, A., Rekhari, S.: Reimagining physiology education with interactive content developed in h5p. *Advances in Physiology Education* **45**(1), 71–76 (2021). <https://doi.org/10.1152/advan.00021.2020>
37. Tran, D., Havlásková, T., Homanová, Z.: Encouraging students to take action in developing problem-solving competency. In: *2019 17th International Conference on Emerging eLearning Technologies and Applications (ICETA)*. pp. 770–776. IEEE (2019). <https://doi.org/10.1109/ICETA48886.2019.9039967>
38. Tripepi, G., Jager, K.J., Dekker, F.W., Zoccali, C.: Selection bias and information bias in clinical research. *Nephron Clinical Practice* **115**(2), c94–c99 (2010). <https://doi.org/https://doi.org/10.1159/000312871>
39. Vallée, A., Blacher, J., Cariou, A., Sorbets, E.: Blended learning compared to traditional learning in medical education: systematic review and meta-analysis. *J Med Internet Res* **22**(8), e16504 (2020). <https://doi.org/10.2196/16504>
40. Van der Merwe, A., Bozalek, V., Ivala, E., Peté, M., Vanker, C., Nagel, L.: Blended learning with technology. In: *Moving beyond the hype: A contextualised view of learning technology in higher education.*, pp. 11–15. Universities South Africa (2015). <https://doi.org/http://hdl.handle.net/2263/50953>
41. Wehling, J., Volkenstein, S., Dazert, S., Wrobel, C., van Ackeren, K., Johannsen, K., Dombrowski, T.: Fast-track flipping: flipped classroom framework development with open-source h5p interactive tools. *BMC Medical Education* **21**(1), 1–10 (2021). <https://doi.org/10.1186/s12909-021-02784-8>
42. Wright, G.B.: Student-centered learning in higher education. *International journal of teaching and learning in higher education* **23**(1), 92–97 (2011)